

Lawrence Berkeley National Laboratory

Recent Work

Title

Complete Genome Sequence of the *Citrobacter freundii* Type Strain.

Permalink

<https://escholarship.org/uc/item/67c4f31w>

Journal

Microbiology resource announcements, 9(19)

ISSN

2576-098X

Authors

Wan, Kenneth H
Park, Soo
Hess, Becky M
et al.

Publication Date

2020-05-01

DOI

10.1128/mra.00240-20

Peer reviewed



Complete Genome Sequence of the *Citrobacter freundii* Type Strain

Kenneth H. Wan,^a Soo Park,^a Becky M. Hess,^b Michael J. Neff,^a Benjamin W. Booth,^a Susan E. Celniker^a

^aBiological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

^bChemical and Biological Signatures, Pacific Northwest National Laboratory, Richland, Washington, USA

ABSTRACT *Citrobacter freundii* is a species of facultative anaerobic Gram-negative bacteria of the family *Enterobacteriaceae*. The complete genome is composed of a single chromosomal circle of 4,957,773 bp with a G+C content of 52%.

Citrobacter freundii is ubiquitous in soil and water and plays an important role in nitrogen fixation in the environment (1, 2). Although it is considered a commensal bacterium of the human intestinal microbiota, it is also an opportunistic pathogen causing infections in humans (3). Importantly, it contains genes that provide multidrug resistance to a broad spectrum of β -lactam antibiotics (4). With limited information regarding the diversity and physiology of *C. freundii*, it was important to sequence the genome of the *Citrobacter freundii* type strain obtained directly from the American Type Culture Collection (strain ATCC 8090), which was originally isolated from canal water (5).

Single colonies were isolated and grown in LB at 37°C. Genomic DNA isolated using Qiagen kits (catalog numbers 19060 and 10262) was sequenced on three platforms, namely, Pacific Biosciences (PacBio), Nanopore, and Illumina platforms. Genomic DNA was sent to GENEWIZ for PacBio library construction and sequencing. Libraries (~10 kb) were constructed using SMRTbell template preparation v1.0, Sequel binding v2.0, and MagBead v2 kits and sequenced on a PacBio Sequel instrument using software v5.0.0.6235 with 10-hour movie times. Default parameters were used for all software tools.

PacBio sequencing yielded a total of 635,329 reads with an N_{50} value of 5,267 bp, totaling 2.6 Gbp. A *de novo* assembly was constructed using Canu v1.7 (6). Sequence quality was assessed using Minimap2 (7). The percentage of reads mapping to the genome is 96.82%, and the estimated coverage is 508 \times .

Nanopore sequencing was performed using the rapid barcoding kit (SQK-RBK004), run with a MinION flow cell (FLO-MIN106D) using MinKNOW v1.13.1, and reads were base called using Guppy v2.0.10. Nanopore sequencing yielded a total of 462,441 reads with an N_{50} value of 8,149 bp, totaling 1.9 Gbp.

Illumina libraries were prepared with the NEBNext Ultra DNA and multiplex oligomer kits (catalog numbers E7370 and E7600) and run on a MiSeq system with v3 chemistry (catalog number MS-102-3003) with paired-end sequencing (2 \times 300 nucleotides). Illumina sequencing yielded a total of 4,421,404 reads passing filter, totaling 1.3 Gbp. The Nanopore and Illumina reads were jointly assembled using Unicycler v0.4.8-beta (8) with 94.16% and 99.95% reads mapping, respectively, and estimated coverages of 358 \times and 269 \times , respectively. Since the assemblies are virtually identical (99.99% average nucleotide identity [ANI]) (9), we analyzed and annotated the Unicycler assembly using RAST v2.0 (10) and NCBI Prokaryotic Genome Annotation Pipeline v4.11 (11). The annotation predicts 4,734 genes in total, with 4,550 protein-coding genes and 119 RNA genes, including 8 rRNA operons, 1 singleton small-subunit rRNA (5S rRNA), 11 noncoding RNAs, and 83 tRNA genes, 1 of which is a tRNA-SeC(p)-TCA. Over 170 genes

Citation Wan KH, Park S, Hess BM, Neff MJ, Booth BW, Celniker SE. 2020. Complete genome sequence of the *Citrobacter freundii* type strain. Microbiol Resour Announc 9:e00240-20. <https://doi.org/10.1128/MRA.00240-20>.

Editor Catherine Putonti, Loyola University Chicago

Copyright © 2020 Wan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Susan E. Celniker, secelnik@lbl.gov.

Received 5 March 2020

Accepted 17 April 2020

Published 7 May 2020

are contained within four candidate prophages (27.7 kb, 32.2 kb, 24.4 kb, and 38.5 kb), representing 2.5% of the genome. As expected from the canal water isolation of this species, KEGG analysis (12) shows that it contains 65 genes that encode and regulate flagellar assembly and function.

Our sequence is similar (ANI, 99.98% [9]) to that of *C. freundii* strain NCTC9750 (GenBank accession number [LR134118](#)) but is smaller by 2,964 bp, differing in the structure of a prophage and insertion of a transposase (InsH) gene. Our sequence also differs in the number of candidate nonessential gene mutations (65 versus 600 in strain NCTC9750).

The identification of tRNA-SeC(p)-TCA led us to search the genome for selenoprotein biosynthesis and metabolism genes. We identified a number of these genes, including *selA*, *selB*, *selD*, *ydf2*, *yedE*, *yedF*, and *ygeY*. At least three proteins likely contain selenocysteine, including those encoded by homologs of the gene for formate dehydrogenase (*fdnG*), a known SeC-containing protein (13). Further characterization of selenoproteins will provide new insights into the roles these proteins play in antioxidant and other activities.

Data availability. The chromosome sequence of *C. freundii* ATCC 8090 is deposited in GenBank under accession number [CP049015](#). The SRA numbers are [SRR11179868](#) and [SRR11179867](#) for the Illumina reads, [SRR11179866](#) and [SRR11179865](#) for the Nanopore reads, and [SRR11179864](#) for the PacBio reads. *C. freundii* strain NCTC9750 was analyzed for comparison.

ACKNOWLEDGMENTS

We thank Alicia Liu and the team at GENEWIZ (South Plainfield, NJ, USA) for library preparation, sequencing reactions, and *de novo* assembly.

This work was supported by the IARPA FELIX program under U.S. Department of Energy contract DE-AC02-05CH11231.

REFERENCES

- French JR, Turner GL, Bradbury JF. 1976. Nitrogen fixation by bacteria from the hindgut of termites. *J Gen Microbiol* 96:202–206. <https://doi.org/10.1099/00221287-95-2-202>.
- Rehr B, Klemme JH. 1989. Formate dependent nitrate and nitrite reduction to ammonia by *Citrobacter freundii* and competition with denitrifying bacteria. *Antonie Van Leeuwenhoek* 56:311–321. <https://doi.org/10.1007/bf00443745>.
- Anderson MT, Mitchell LA, Zhao L, Mobley H. 2018. *Citrobacter freundii* fitness during bloodstream infection. *Sci Rep* 8:11792. <https://doi.org/10.1038/s41598-018-30196-0>.
- Pepperell C, Kus JV, Gardam MA, Humar A, Burrows LL. 2002. Low-virulence *Citrobacter* species encode resistance to multiple antimicrobials. *Antimicrob Agents Chemother* 46:3555–3560. <https://doi.org/10.1128/aac.46.11.3555-3560.2002>.
- Werkman CH, Gillen GF. 1932. Bacteria producing trimethylene glycol. *J Bacteriol* 23:167–182. <https://doi.org/10.1128/JB.23.2.167-182.1932>.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Luo C, Rodriguez RL, Konstantinidis KT. 2013. A user's guide to quantitative and comparative analysis of metagenomic datasets. *Methods Enzymol* 531:525–547. <https://doi.org/10.1016/B978-0-12-407863-5.00023-X>.
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufu S, Li W. 2013. Prokaryotic Genome Annotation Pipeline. In *The NCBI handbook*, 2nd ed. National Center for Biotechnology Information, Bethesda, MD. <https://www.ncbi.nlm.nih.gov/books/NBK174280>.
- Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35:W182–W185. <https://doi.org/10.1093/nar/gkm321>.
- Berg BL, Baron C, Stewart V. 1991. Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. II. Evidence that a mRNA stem-loop structure is essential for decoding opal (UGA) as selenocysteine. *J Biol Chem* 266:22386–22391.